

Standards Validation

April 23, 2008

Alaska Department of
Early Education and Development

Dillard Research Associates

Welcome and Thank You

- Introductions of Participants
- Purpose and Agenda
- Overview of Big Ideas in Three Parts
 - Training on Scoring Protocols and Student Materials
 - Comparison of Items and Tasks
 - Judgments of Proficiency
- Explanation of statistical studies

- Comparisons between 2007 and 2008
 - 2007 is not grade banded
 - Store bought test (as is)
 - More items per tasks
 - Wider range of tasks
 - 2008 is grade banded and divided into strands and exactly match the EXGLS
- Some items are in common (called linking items) between 2007 and 2008

- Step 1. Train on Test Administration, attending to both scoring protocols and student materials.
- Step 2. Review Extended Grade Level Expectations
- Step 3. Review cross walk document and relevant tasks; consider other tasks to grade level band
- Step 4: Review draft proficiency level descriptors (PLDs)
- Step 5. Establish proficiency levels for each task and for each level
- Step 6. Confer with partner and articulate your rationale. Consider the issues in item 1 above).
- Step 7. Review an example data set to confirm your reasoning of each task.
- Step 8. As a group, devise a combined (across task) judgment for each proficiency level.
- Step 9. Complete the Standard Setting Form for your grade level.
- Step 10. Review Impact Data

Comparison of Items and Tasks

- Linking items are embedded in some tasks
- The items are either exact or virtual in their comparability
- Point totals are different (be careful)
- Two booklets are used: (a) Items and Tasks, and (b) Judgment Responses
- Two responses are needed: (a) rating of difficulty similarity and (b) comments
- Summary Judgment is needed at the end of each grade band.

Comparison of Items and Tasks

- Linking items are described
 - Task-skill construct
 - 07 Item # and 08 Item #
 - Item description
- Linking items are compared statistically
 - N = Count of students who took the item
 - Average = Mean of the sample
 - SD = Standard deviation (average amount of variation around the mean)

07 N	07 Ave	07 SD	08 N	08 Ave	08 SD
106	1.72	0.69	116	1.44	0.90
105	1.24	0.97	103	1.24	0.96

Person Centered Judgments of Proficiency

- Student Profiles
 - Sampled 1 student separating Proficient from Below
 - Sampled 2 students Below
 - Sampled 2 students Above
 - Sampled 1 student separating Proficient from Advanced
 - Sampled 2 students Below
 - Sampled 2 students Above
 -

Person Centered Judgments of Proficiency

- Students randomly distributed within the grade band
- Tasks are noted with each performance value
- Pattern of values is very different among the group: TAKE NOTE OF THIS VARIATION.
- Categorize the student as P(roficient) or B(elow)
- Review the cut scores from 2007 to finalize your judgment

Person Centered Judgments of Proficiency

- Review the predicted categories from an initial run of the data
 - 2007 includes only Below and Above
 - 2008 includes
 - Advanced, Proficient, Below, Far Below
 - Above and Below
- Go back over you judgments and revise them in you feel they are in error.

Summary of Process

- Review of judgments and adjustment of ratings/comments
- Consideration of outcomes
- Open mic on issues of process
- Evaluation of workshop